FIG.2A

```
LOCUS         HSU22027      7215 bp      DNA            PRI      22-OCT-1995
DEFINITION    Human cytochrome P450 (CYP2A6V2) gene, complete cds.
ACCESSION     U22027
NID           g1008461
KEYWORDS      .
SOURCE        human.
  ORGANISM    Homo sapiens
              Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
              Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE     1  (bases 1 to 7215)
  AUTHORS     Fernandez-Salguero,P., Hoffman,S.M., Cholerton,S.,Mohrenweiser,H.,
              Raunio,H., Rautio,A., Pelkonen,O., Huang,J.D., Evans,W.E.,
              Idle,J.R. et.al.
  TITLE       A genetic polymorphism in coumarin 7-hydrozylation: sequence of the
              human CYP2A genes and identification of variant CYP2A6 alleles
  JOURNAL     Am. J. Hum. Genet. 57 (3), 651-660 (1995)
  MEDLINE     95397851
REFERENCE     2  (bases 1 to 7215)
  AUTHORS     Fernandez-Salguero,P.
  TITLE       Direct Submission
  JOURNAL     Submitted (01-MAR-1995) Pedro Fernandez-Salguero, National
              Institutes of Health, 9000 Rockville Pike, Bethesda, MD 20894, USA
FEATURES             Location/Qualifiers
     source          1..7215
                     /organism="Homo sapiens"
```

FIG.2A CONT.

```
5'UTR    782..790
CDS      join (791..970, 1237..1399, 2115..2264, 2499..2659,
         3207..3383, 4257..4398,4873..5060,5577..5718, 6308..6489)
         /gene=CYP2A6V2:
         /codon_start=1
         /product=cytochrome P450"
         /db_xref=PID:g1008462"
         /translation=MLASGMLLVALLACLTVMVLMSVWQQRKSKGKLPPGPTPLPFIG
         NYLQLNTEQMYNSLMKISERYGPVFTIHLGPRRVVVLCGHDAVREALVDQAEEFSGRG
         EQATFDWVFKGYGVVFSNGERAKQLLRFAIATLRDFGVGKRGIEERIQEESGFLIEAI
         RSTHGANIDPTFLSRTVSNVISSIVFGDRFDYKDKEFLSLLRMLGIFQFTSTSTGQ
         LYEMFSSVMKHLPGPQQQAFQLLQGLEDFIAKKVEHNQRTLDPNSPRDFIDSFLIRMQ
         EEEKNPNTEFYLKNLMMSTLNLFIAGTETVSTTLGYGFLLMKHPEVEAKVHEEIDRV
         IGKNRQPKFEDRAKMPYMEAVIHEIQRFGDVIPMSLARRVKKDTKFRDFFLPKGIEVF
         PMLGSVLRDLRFFSNPRDFNPQHFLGEKGQFKKRDAFVPFSIRKRNCFGEGLARMELF
         LFFTTVMQNFRLKSSQSPKDIDVSPKHVGFATIPRNYTMSFLPR

exon     791..970
         /gene=CYP2A6V2:
         /number=1

exon     1237..1399
         /gene=CYP2A6V2:
         /number=2

exon     2115..2264
```

3/59

FIG.2A CONT.

```
exon            /gene=CYP2A6V2:
                /number=3
                2499..2659
                /gene=CYP2A6V2:
                /number=4
                3207..3383
exon            /gene=CYP2A6V2:
                /number=5
                4256..4398
exon            /gene=CYP2A6V2:
                /number=6
                4873..5060
exon            /gene=CYP2A6V2:
                /number=7
                5577..5718
exon            /gene=CYP2A6V2:
                /number=8
                6308..6489
exon            /gene=CYP2A6V2:
                /number=9
                6490..6744
3'UTR
BASE COUNT    1646 a     2196 c     1746 g     1627 t
ORIGIN
```

FIG.2A CONT.

BASE COUNT

```
   1 aagttcccct gaaatatggc tctggtcttc ctccccttgc caatgaagaa gatggcagtg
  61 gaggttctat ggcagccatc ctggcctcac tctgaggttc caatgaggat tctgggcatc
 121 aagagacagc tctgggcaaa gctaaatcaa gtcagcccct ggaccagtg ctgggctgct
 181 gggcttctg ggagaacgcc gctgggcttg ctacacactc ctcctcccag aaactccaca
 241 cccacagccc tgggtcttcc tagcccgag actttcaagt ccatatgcct ggaatccccc
 301 ttcctgagac cctaaccct gcatcctcca caacagaaga ccctaaatg cacagccaca
 361 ctttgtctta ccctaataaa accagacct ttggattcct ctccctgga accccagat
 421 ccgcacaact ttggggtgca ttctcactct cagaccccaa atccaaagcc caagtgctcc
 481 cctatgcaaa tattccaaac tcctcagttc tacagcttat ctgttgcccc ctcctaaatc
 541 cacagccctg cggcaccct cctgaagtac cacagattta gtctggaggc ccctctctg
 601 ttcagctgcc ctgggtccc cttatcctcc gtgctggct gtgtccaaag ctaggcagga
 661 ttcatggtgg ggcatgtagt tggggagtga aatgaggtaa ttatgtaatc agccaaagtc
 721 catccctctt tttcaggcag tataaggca aaccacccca gccgtcacca tctatcatcc
 781 ctctaccacc atgctggcct cagggatgct tctgctggcct gcctgactgt
 841 gatggtcttg atgtctgttt ggcagcagag gaagacagca gggaaagctgc ctcggggacc
 901 caccccattg cccttcattg gaaactacct acagagcaga tgtacaactc
 961 cctcatgaag gtgtcccaag acaggggagat gggtgtctcg gggtgggggc tgcctagttg
1021 gctggggctt tgtggcaggg ggttgaccag agtcttagga aatggagttt
1081 tggagtttca gcatcagaaa gacaggatct tgggaccag agctccctga ctgtgagaac
1141 ctgggtgcga agcatcccag cacatgacat ctcggtgctg ggccccattc agagtggagg
1201 gttctccctc taaccactcc caccccactc catcagatca gtgagcgcta tggcccccgtg
1261 ttcaccattc acttggggcc ccggcggggtc gtggtgctgt gtggacatga tgccgtcagg
```

FIG.2A CONT.

```
1321  gaggctctgg tggaccaagc tgaggagttc agcggggcgag gcgagcaagc caccttcgac
1381  tgggtcttca aaggctatgg tgcccaagag ggggaaggtg ggcaggtgga cacgaaggtc
1441  tcagtgttcc cagccttctc cctgactctc ctgacaactg gaggataagg gagagtcccc
1501  agctctggtct tcccctccca tctcccctaca ttggggccctc tccatgtgta tccctcacct
1561  gtctccaagcg gccctgtcct gattcctctct tgcctctctc tgcccccacct ccttattctc
1621  tctcactgga gtcctctctt tccccctct tgccctct ctccatctct tgggtttctg
1681  tttaccaagcc ctggggtctt ctcatctctc cggatcccttt tctcttatct taggatgcca
1741  ctctgggtttt ctcatctctc cggatcccttt tctcttgtct tctctctcc ttctggctt
1801  gggttattcc tacttcaca tctttcaggct ccaatctctcc gtaacagttc ctcttccttc
1861  cagacccttct ctgtttctatt ctcaatatta aactctctgc tccagtcag cttaagaatc
1921  tcaccaagg agaggatgtc ctcacaccag atctcccat atctccctcc ccacccctcc
1981  atcctctgcc tccatcactc tctttctctc cccactgccg atcctggacgcg atccaatgga
2041  gtgtggagct aatgccgtga agctatgtgc atctctctgt ctggccgtac ctgggtaata
2101  acctgatcga ctaggccgtgg tattcagcaa cgggagcgc gccaagcaag tcctcgctt
2161  tgccatcgcc acctcgaggg gggcaagcga ggcaagcagg agcgcatcca
2221  gggaggagtcg ggcttcctca tcgaggccat cggagcacg cacggtgagc agggagaccca
2281  gagtgcgggg gcaggagaag gaaaacaccc aggacgagga accgcgcgc gttctgcctg
2341  gggatgggga ctaggtgggg aaaggcgccc gcacttccag ccctggagtc tgccgctggg
2401  aatttggctc aacaaggccc aattctgact ctcctcagac ctctgagttg
2461  actctctcca caaccccctt ctcccgacat accggaggc gccaatatcg atcccaccctt
2521  cttcctgagc cgcacaagtc caatgtcat cagctccatt gtctttgggg accgcttga
2581  ctataaggac aaagagttcc tgtcactgtt gcgccatgatg ctaggaatct tccagttcac
2641  gtcaacctcc acggggcagg taatggttgc agcccggccc gtgaaggccc ttaccaaaaac
```

FIG.2A CONT.

```
2701 cggcaaattg ttcccctacc gggggaaggg ggccccaaat tccaccgcc ccccggacag
2761 tgtcccctca aaatcagtcc ccgatttggg caaattggca gagtggaacc agaccgggt
2821 tggttgtcca atcccctgct ctccagggac accggatag cacaacagat gctcccaaa
2881 acagagcctg ctggcaggat gcatacctc agctcagctc tctcaccctg ggcacgtgtt
2941 cccatcccca acttaccggt aatttctaac agatgctccc taccaggtc ttcttgaata
3001 ttttaacacc cggaaaccct gggtacctaa ccttcctgt aaacttaga gattagttcc
3061 tatccggccc ctctgaaata cctaaccacc ggagaccaga tgcctttaac tcagttcctt
3121 ccttgctatg aaacaaatcc cattcccatc agctcctgcc ccgtgacagc tgtccttccc
3181 ttcccatcct ctctctgcaa cccagctct atgagatgtt ctcttcggtg atgaaacacc
3241 tgccaggacc gcagcaacag gcctttcagt tgctgcaagg gctggaggac ttcatagcca
3301 agaaggtgga gcacaaccag cgcacgctgg atcccaattc cccacgggac ttcattgact
3361 cctttctcat ccgcatgcag gaggtacacc ccagcagcca ctgcgggggag atgcaaaagcc
3421 aggcagaggg aaatcagtct gggagtgggg caggcagatg acacaggccc attcaaatta
3481 accctcatca taataatcct cacaattggc tggttgccgt ggctaacagc ctgtaatccc
3541 agcactttgg gaggccgagg caggtggatc acctgaggtc aggagttcga gaccagcctg
3601 gccaacatgg tcaaacccg tctctactaa aaatacaaaa attagttggg catggtggcg
3661 cgaagggggg cagaggttgc aatgagccaa gatcacgcca ttgcactcca gtctgggtga
3721 cagaatgagg ccctgtgtca aaaaaaaata atcacttgtt taaaaagtaa gtgagcctgc
3781 atggtcatgc gcatgtgcag ctccagctac tcaggaggct gaggcaggag gattgcttga
3841 gctcaggagt tggcgtccgg cctgtgcaac ttagcaagac caagtcagta taagaaaaaa
3901 aaaaaacaaa aaaaaagctg acagctaagt tgataattga cggacagatg gtcagcaagg
3961 taacgaaggt gagaaggaag agcattgggg gcaacgccag gagtcaagggc aagggctggt
```

FIG.2A CONT.

```
4021 tcctagagcg agtctggtag gatctagggc ccctcttctc caccctgcgg tcttgcccca
4081 aagagaggtc gagggtgctg ggattgcgct agactcgagt ctgtgtagat cttggggtcc
4141 cctcttgacc cccattggtc tgaacctaag agtggaagat ccatgggtg aaccctaga
4201 tggtgccctg aggtcaagca ggagtgaggt tgtcctaaag cccctctcc cttcaggagg
4261 agaagaaccc caaacaggag ttctacttga agaacctgat gatgagcacg ttgaacctct
4321 tcattgcagg caccgagacg gtcagcacca ccctgcacta tggcttctta ctgctcatga
4381 agcacccaga ggtggaaggt aaggctggaa gggggacgaa gtggagggc ccagaccctc
4441 aaaattcccc ttcgactggt gcaatgtccc cacctgtccc agatcccggg accctggagac
4501 gtgactggct gtccagagac agggcaacat tcagctggta ggcatcaagc gagtctcatt
4561 agatattaaa atattgaaaa tgtctgcact gattggtcag tcacttctgt cccaagccca
4621 ctgagtgccc actgcccgtt ccacgggtc atcccctaag ttcctcccct tgcctccct
4681 gtgattctgg cacaacctgg ttaacaggat cctactccaa caatgcgaat gggtgatgtc
4741 tgttctgtta tgaatgctct acttccgtct cataggcgga ggcatttcat ccacccatt
4801 ttgcctatcc ggactatcat ttcctgctct gagaccccta gatacctaaa cacattcccc
4861 ctcctccccc agccaaggtc catgaggaga ttgacagagt gatcggcaag aaccggcagc
4921 ccaagtttga ggaccgggc aagatgccct acatggaggc agtgatccac gagatccaaa
4981 gattggaga cgtgatcccg cctgatccag atgagtttgg cccgcagagt caaaaaggac
5041 gggatttctt cctccctaag gtgctatccg ccccacccc ccagactcag gggactccag
5101 cccctctctg tgtcccagc atcccacaca cattagaagc tttctccact ctgtcccact
5161 cccctcaatca gtcaaaaaag acttcccaa ccaccacatc cgttccacct ttccacttag
5221 acactcctga gtcctgcatc tctccagact cttgtgtca ggagaatcaa acacatgttc
5281 ccaaacttcc tatcttaaga aacagaagc ccctttccat tcggcccttt gtcatagga
5341 cagaaatctc aggtcccccca aactcctgcc tagaaggaca tggacccat gtctcccaaa
```

FIG.2A CONT.

```
5401 cttcctgtt  cagagatgtg  aaccttctat  ccccaaggt   cctccctcag  aggtccccaa
5461 ttccatgcc  tgccacttcc  cctcaccggg  gcaccctagt  tccccctcca  gcccctgt
5521 actctcaaca atccccaaac  ccgcctcatc  acatacacct  tcctcctccc  tccagggca
5581 tagaagtgtt ccctatgttg  ggctccgtgc  tgagagacct  caggttcttc  tccaacccc
5641 gggacttcaa tcccagcac   ttcctgggtg  agaaggggca  gtttaagaag  cgtgatgctt
5701 ttgtgccctt ctccatcagt  aagagaccac  tgtttggtgc  caggcttact  actcacaca
5761 gcaggggcct ccctaccca   gttcccctct  ctgccgtgta  gcctagtatt  tcccagctt
5821 ggcaagttcc tgttagcaat  ctaccgtcga  gcaccaaggt  gatactcctt  taactaccaa
5881 gcaccagta  cctgtgccca  aaggaaaca   caatcccctt  tcccagcgg
5941 gggaaaaca  aaggccagag  agaatcagag  atttattcc   ctaggtcac   acaggagatt
6001 cttcagcatc cctaaaaagg  agatgacggc  acagcaggtc  atatttggga  gttcttatct
6061 gggggaaggg ggatcttaaa  cctccattg   tggacaacctg gcatcgatca  accccatctt
6121 ttggtcatct tttgggtcac  tcaaggaaac  tgaggtcaag  gaggtcccctc
6181 ttaaagtctc tcaaggccat  atattccacc  cttcctcct   gggagagaccg cagctggag
6241 tcggtactgg ggcgaggctg  cactgagagt  gggcttcacc  tccaccctc   cgccctctcc
6301 tcctcaggaa agcggaactg  tttcggagaa  ggcctggcca  gaatggagct  ctttctcttc
6361 ttcacacacg tcatgcagaa  cttcccgcctc aagtcccctc  agtcacttga  gacacattgac
6421 gtgtccccca aacacgtggg  ctttgccacg  atcccacgaa  actacacaat  gagcttcctg
6481 cccgctgag  cgaggctgt   gccggtgggg  cgggccaggt  cgggggcagg  gaaaggcag
6541 ggccaagacc gggcttggga  gaggggcgca  gctaagactg  ggcgcagggat ggcggaaag
6601 aaggggcgtg gtggctagag  ggaggagaag  aaacagaagc  ggctcagttc  accttgataa
6661 ggtgcttccg agctgggatg  agaggaagga  aacccttaca  ttatgctatg  aagagtagta
```

9/59

FIG.2B

LOCUS   HSP452B6   1415 bp   RNA   PRI   29-MAY-1992
DEFINITION   Human MRNA FOR CYTOCHROME P-450IIVB6.
ACCESSION   X13494
NID   g35206
KEYWORDS   Cytochrome; cytochrome P4501IB6.
SOURCE   human.
  ORGANISM   Homo sapiens
             Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
             Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1415)
  AUTHORS   Miles, J.S.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-NOV-1988) Miles J.S., Imperial Cancer Research Fund,
             Lab of Molecular Phrmacology and Drug Metabolism, Hugh Robson
             Building, George Square, Edinburgh, EH8 9XD
REFERENCE   2 (bases 1 to 1415)
  AUTHORS   Miles, J.S., McLaren, A.Q. and Wolf, C.R.
  TITLE   Alternative splicing in the human cytochrome P450IIB6 gene
             generates a high level of aberrant messages
  JOURNAL   Nucleic Acids Res. 17 (20), 8241-8255 (1989)
  MEDLINE   90045947
COMMENT   The sequence is a compilation of genomic and cDNA clones. **map:
             chromosomal location=19q12-13.2;
             Data kindly reviewed (13-NOV-1989) by Miles, J.S.
FEATURES   Location/Qualifiers

FIG.2B CONT.

```
source        1..1415
              /organism="Homo sapiens"
misc-feature  9..110
              /note=exon 1, partial"
misc-feature  111..273
              /note=exon 2"
misc-feature  274..423
              /note=exon 3"
misc-feature  424..584
              /note=exon 4"
misc-feature  585..761
              /note=exon 5"
misc-feature  762..903
              /note=exon 6"
misc-feature  904..1091
              /note=exon 7"
misc-feature  1092..1233
              /note=exon 8"
misc-feature  1234..1415
              /note=exon 9", coding region"
BASE COUNT    341 a    430 c    328 g    316 t
ORIGIN
    1 gaattccgcc ctgcaccat gaccgcctcc caccagggcc ccgcccctctg cccctttgg
```